Probing the qudit depolarizing channel

Michael Frey

Department of Mathematics, Bucknell University, Lewisburg, PA 17837, USA∗

David Collins

Department of Physical and Environmental Sciences, Mesa State College,
1100 North Avenue, Grand Junction, CO 81501, USA†

Karl Gerlach

ITT Corp., 2560 Huntington Avenue, Alexandria, VA 22303, USA

For the quantum depolarizing channel with any finite dimension, we compare three schemes for channel identification: unentangled probes, probes maximally entangled with an external ancilla, and maximally entangled probe pairs. This comparison includes cases where the ancilla is itself depolarizing and where the probe is circulated back through the channel before measurement. Compared on the basis of (quantum Fisher) information gained per channel use, we find broadly that entanglement with an ancilla dominates the other two schemes, but only if entanglement is cheap relative to the cost per channel use and only if the external ancilla is well shielded from depolarization. We arrive at these results by a relatively simple analytical means. A separate, more complicated analysis for partially entangled probes shows for the qudit depolarizing channel that any amount of probe entanglement is advantageous and that the greatest advantage comes with maximal entanglement.

PACS numbers: 03.67.-a

I. INTRODUCTION

A quantum operation, or channel, Γ is a trace-preserving, completely positive map for representing the change of state of a quantum system. Quantum channels are the fundamental building blocks of quantum information processing. In any physical implementation, though, the channel is usually not fully known and must be determined experimentally by observing its effect on prepared quantum systems, or probes. This is the problem of quantum channel identification [1–7], also known as quantum process tomography [8], and a major challenge is to determine the optimal quantum state and measurement for the channel probe. In particular, the probe may be entangled with an ancilla system, or with other probes.

Quantum channel identification is statistical by nature and is generally formulated as a parameter estimation problem: the unknown channel Γ_θ is supposed to belong to a parametric family $\{\Gamma_\theta, \theta \in \Theta\}$ of channels, and we identify the channel by estimating θ. The general scheme for doing this is to prepare the channel input probe in a chosen quantum state σ, make a quantum measurement $\mathcal{M}$ of the channel output $\rho_\theta = \Gamma_\theta[\sigma]$, and base the estimation of θ on the measurement's registered result X. This process (see Fig. 1) is repeated to obtain n independent, identically distributed measurement outcomes $X_1, \ldots, X_n$, and then θ is estimated by an estimator $\hat{\theta} = \hat{\theta}(X_1, \ldots, X_n)$. This arrangement, in which

a prepared quantum system is used to probe a quantum channel, is standard for channel identification, it is common in interferometry [9], and it occurs generally whenever a quantum measurement is made indirectly by observing a change in an ancillary part of the measurement apparatus.

$$\sigma \rightarrow \boxed{\Gamma_\theta} \rightarrow \rho_\theta \rightarrow \boxed{\mathcal{M}} \rightarrow X \xrightarrow{n \text{ runs}} \hat{\theta} = \hat{\theta}(X_1, \ldots, X_n)$$

FIG. 1. Set-up for probing a depolarizing channel Γ_θ by a probe system prepared in state σ.

We study the problem of identifying the qudit (dimension $d < \infty$) depolarizing channel. This channel is a standard model of quantum noise, it is analytically tractable, and it is the basis of investigation in a range of contexts [7, 10–14]. Our study of the qudit depolarizing channel proceeds most directly from the works of Fujiwara and Sasaki *et al.* Fujiwara [1] considers the qubit ($d = 2$) depolarizing channel with probes prepared by three different schemes. Using a likelihood approach based around quantum Fisher information, Fujiwara obtains comprehensive "local/asymptotic" optimality results for his three identification schemes, even including probes prepared in partially entangled states. Sasaki *et al.* [2] study the more general qudit ($d < \infty$) depolarizing channel from a Bayesian perspective, comparing the same channel identification schemes as Fujiwara. Their Bayesian approach yields "global" optimality results for any $d < \infty$, though their results for partial entanglement are limited to $d = 2$. We take the same approach as Fujiwara but treat the general qudit depolarizing channel, introducing a common simple expression for the quan-

---

∗ mfrey@bucknell.edu; Author to whom correspondence should be addressed.
† dacollin@mesastate.edu

tum Fisher information to address all three identification schemes treated in [1, 2]. We extend this method to entanglement with an ancilla that is itself depolarizing and to a scheme we call probe re-circulation, whereby the probe is passed multiple times through the channel before measurement. By a separate analysis we show for the depolarizing channel that any degree of entanglement is beneficial for qudits of any dimension and that maximal entanglement is most beneficial. This broad usefulness of entanglement in any dimension is known in only a few instances [15].

The remainder of the paper is arranged as follows. We set the stage for our work in Sec. II by recalling the definition of quantum Fisher information and its role in the quantum Cramér-Rao inequality. We then derive a general expression for the quantum Fisher information in the output of the qudit depolarizing channel. In Sec. III we apply this expression to channel identification schemes in which the qudit probe is unentangled, maximally entangled with an ancilla qudit, or maximally entangled with another probe (the three schemes studied in [1, 2]). For each of these schemes, we treat the possibility that the probe is circulated back through the channel before measurement. We further apply our expression for quantum Fisher information to a probe that is maximally entangled with an ancilla qudit that is itself depolarizing. Partially entangled probes must be addressed with a separate method; we do this in Sec. IV. We take up the question of optimal probe measurement in Sec. V and summarize our results in Sec. VI.

II. PRELIMINARIES

The quantum Fisher information $J(\theta) = J[\rho_\theta]$ in the parametric output $\rho_\theta = \Gamma_\theta(\sigma)$ of a quantum channel bounds the ultimate precision of the estimation of the parameter θ attainable by quantum measurement $\mathcal{M}$ of ρ_θ. The quantum information (Cramér-Rao) inequality states for n independent, identically distributed registrations of any quantum measurement $\mathcal{M}$ and any unbiased estimator $\hat{\theta}$ that

$$\text{mse}\,[\hat{\theta}] \geqslant \frac{1}{nJ(\theta)} \tag{1}$$

where $\text{mse}\,[\hat{\theta}] = \text{E}\,[(\hat{\theta} - \theta)^2]$ is the mean squared error of $\hat{\theta}$. The quantum Fisher information $J(\theta)$ in Eq. (1) is

$$J(\theta) = J[\rho_\theta] = \text{tr}[\rho_\theta L_\theta^2] \tag{2}$$

where L_θ is the quantum score operator (symmetric logarithmic derivative) of ρ_θ defined by

$$\frac{L_\theta \rho_\theta + \rho_\theta L_\theta}{2} = \partial_\theta \rho_\theta \tag{3}$$

where ∂_θ signifies differentiation. Quantum Fisher information originated with Helstrom [16]; its relationship to

classical Fisher information is brought out in [17]; and recent presentations are [18, 19]. The Cramér-Rao bound, Eq. (1), is asymptotically achievable so the larger $J(\theta)$ is in Eq. (1), the more precisely θ can potentially be estimated. We therefore interpret $J(\theta)$ to be the overall information about θ available in a parametrically defined quantum state ρ_θ. In this sense $J(\theta)$ serves as a quantitative measure of the relative merit of different channel identification schemes. The information inequality, Eq. (1), and this interpretation of quantum Fisher information, readily extends [18] to biased estimators, dependent and non-identically distributed registrations, vectors of channel parameters [20], and transformations of parameter.

The qudit (d-dimensional) depolarizing channel defined for any qudit input state σ is

$$\Gamma_\theta(\sigma) = \frac{1-\theta}{d}\mathbf{I}_d + \theta\sigma \tag{4}$$

where $1 - \theta$ is the probability of depolarization and $\frac{1}{d}\mathbf{I}_d$ is the completely mixed qudit state. The channel Γ_θ is completely positive for $0 \leqslant \theta \leqslant 1$. In fact [21], Γ_θ is completely positive for $-1/(d^2 - 1) \leqslant \theta \leqslant 1$; we restrict our attention to $0 \leqslant \theta \leqslant 1$ where θ reflects solely depolarization. We take advantage of a general transformation of parameter to simply and flexibly express the quantum Fisher information $J(\theta)$ for the qudit depolarizing channel. This formulation of $J(\theta)$ is used in the following section to develop and compare schemes for estimating θ and identifying Γ_θ. In particular, we generalize results known [1] for the qubit depolarizing channel to the setting of qudits.

We will need to consider a variety of depolarizing-type channels, whose general form is

$$\sigma \rightarrow \frac{1 - h(\theta)}{d}I_d + h(\theta)\sigma \tag{5}$$

where d is the dimension of the system and $0 \leqslant h(\theta) \leqslant 1$ is monotonic and has the interpretation that $1 - h(\theta)$ is a generalized depolarizing probability. Given that the probe state σ at the input of the qudit depolarizing channel has spectral decomposition $\sigma = \sum_{i=1}^{d} s_i |s_i\rangle\langle s_i|$ it follows that $\rho_\theta = \Gamma_\theta(\sigma) = \sum_{i=1}^{d} \lambda_i(\theta)|s_i\rangle\langle s_i|$ where

$$\lambda_i(\theta) = \frac{1 - h(\theta)}{d} + s_i h(\theta). \tag{6}$$

Only the eigenvalues $\lambda_i(\theta)$ of ρ_θ depend on θ; its eigenvectors $|s_i\rangle$ do not, so ρ_θ is quasi-classical [18, 20]. In this situation ρ_θ and the score operator L_θ commute, L_θ is readily calculated, and we have

$$J(\theta) = \text{Tr}\,[\rho_\theta L_\theta^2] = \text{Tr}\,[\rho_\theta^{-1}(\partial_\theta \rho_\theta)^2] = \sum_{i=1}^{d} \frac{\lambda_i'(\theta)^2}{\lambda_i(\theta)}. \tag{7}$$

Fujiwara [1] observed that the quantum Fisher information $J(\theta) = J[\Gamma_\theta(\sigma)]$ in a channel output $\Gamma_\mu(\sigma)$ about

a channel parameter θ is maximized by a pure state σ. This observation applies in particular to the qudit depolarizing channel, so we restrict our attention here to pure probe states. Let $\{|s_i\rangle, i = 1, \ldots, d\}$ be an orthonormal basis relative to which $\sigma = |s_1\rangle\langle s_1|$. The channel output is then

$$\rho_\theta = \frac{1 + (d-1)h(\theta)}{d}|s_1\rangle\langle s_1| + \frac{1 - h(\theta)}{d}\sum_{i=2}^{d}|s_i\rangle\langle s_i|. \quad (8)$$

Reading the eigenvalues of ρ_θ from Eq. (8), we find that the quantum Fisher information Eq. (7) in ρ_θ is

$$J(\theta) = \frac{(h'(\theta))^2}{(1 - h(\theta))\left(h(\theta) + \frac{1}{d-1}\right)}. \quad (9)$$

This expression for $J(\theta)$ applies to all the qudit depolarizing channel identification schemes of the next section.

III. PROBING WITH ENTANGLEMENT

We first consider channel identification in which the qudit probe is prepared in an unentangled pure state σ—we call this scheme O. The probe is applied to the depolarizing channel with depolarizing probability $1 - \theta$; i.e., $h(\theta) = \theta$. We have immediately from Eq. (9) that the quantum Fisher information available in the channel output about θ is

$$J_{\mathrm{O}}(\theta) = \frac{1}{(1 - \theta)\left(\theta + \frac{1}{d-1}\right)} \quad (10)$$

with no input entanglement. This quantum Fisher information is the same for any pure input state σ, because the qudit depolarizing channel is unitarily invariant and $J[\rho_\theta] = J[\mathbf{U}\rho_\theta\mathbf{U}^\dagger]$ for any unitary transformation $\mathbf{U}$.

We now consider an externally entangled probe. Specifically, in scheme E one of a maximally entangled pair of qudits is passed through Γ_θ with depolarizing probability $1 - \theta$; again $h(\theta) = \theta$. The state of the maximally entangled qudit pair is $\sigma = |\mu\rangle\langle\mu|$ with

$$|\mu\rangle = \frac{|s_1 t_1\rangle + \cdots + |s_d t_d\rangle}{\sqrt{d}} \quad (11)$$

where $\{|s_i\rangle\}$ and $\{|t_i\rangle\}$ are orthonormal bases for each qubit. The channel output $\rho_\theta = (\mathbf{I} \otimes \Gamma_\theta)(\sigma)$ is, after some calculation,

$$\rho_\theta = \frac{1 - \theta}{d^2}\mathbf{I}_{d^2} + \theta\sigma . \quad (12)$$

This output is identically that of a d^2-dimensional qudit depolarizing channel with depolarizing probability $1 - \theta$ and a probe prepared in a pure state. From Eq. (10), then, the quantum Fisher information in Eq. (12) about θ obtained by a maximally entangled probe is

$$J_{\mathrm{E}}(\theta) = \frac{1}{(1 - \theta)\left(\theta + \frac{1}{d^2-1}\right)}, \quad (13)$$

and the information gained using a maximally entangled probe is

$$J_{\mathrm{E}}(\theta) - J_{\mathrm{O}}(\theta) = \frac{d(d-1)(1-\theta)}{(1 - \theta + d\theta)(1 - \theta + d^2\theta)}. \quad (14)$$

This gain is positive for all $d < \infty$ and all $\theta < 1$.

One might entangle two qudits and use *both* as probes as a scheme for channel identification. In scheme B two qudits in the maximally entangled state $\sigma = |\mu\rangle\langle\mu|$, with $|\mu\rangle$ as in Eq. (11), are each passed through the qudit depolarizing channel. This yields the channel output $\rho_\theta = (\Gamma_\theta \otimes \Gamma_\theta)(\sigma)$, in which case

$$\rho_\theta = \frac{1 - \theta^2}{d^2}\mathbf{I}_{d^2} + \theta^2\sigma. \quad (15)$$

This is identically the parametric family of output states that would result from passing a d^2-dimensional unentangled qudit in the pure state σ through a channel with depolarizing probability $1 - h(\theta)$ where $h(\theta) = \theta^2$. So from Eq. (9), the information *per channel use* in Eq. (15) about θ is

$$J_{\mathrm{B}}(\theta) = \frac{2\theta^2}{(1 - \theta^2)\left(\theta^2 + \frac{1}{d^2-1}\right)}. \quad (16)$$

Our three channel identification schemes O, E, and B are just those considered in [1, 2]. These schemes can be compared in different ways based on the quantum Fisher informations $J_{\mathrm{O}}(\theta)$, $J_{\mathrm{E}}(\theta)$ and $J_{\mathrm{B}}(\theta)$. The three schemes use different resources to different degrees, suggesting different accountings. Schemes O and E involve just one channel use; scheme B requires two uses. Schemes E and B use entanglement, whereas scheme O does not. Schemes E and B involve a dimensional extension of the channel (with a more complex measurement); scheme O does not. Both Fujiwara [1] and Sasaki *et al.* [2], compared these schemes on the basis of channel extension. Our $J_{\mathrm{O}}(\theta)$, $J_{\mathrm{E}}(\theta)$, and $J_{\mathrm{B}}(\theta)$ are information gained per channel use. Thus Fujiwara effectively compared $2J_{\mathrm{O}}(\theta)$, $J_{\mathrm{E}}(\theta)$ and $2J_{\mathrm{B}}(\theta)$. He found by his accounting that no one scheme uniformly dominated the others; each scheme was optimal for different θ.

We compare our identification schemes O, E, and B for the general qudit depolarizing channel on the basis of quantum Fisher information per channel use. This comparison emphasizes the essential cost of the channel to estimation. From a resource perspective this seems most natural, because if channel use is not relatively costly, we would always just probe the channel by scheme O with as many probes as needed to achieve the desired precision. Comparing quantum Fisher information per channel use, we find that

$$J_{\mathrm{E}}(\theta) \geqslant J_{\mathrm{O}}(\theta) \quad \text{and} \quad (17a)$$
$$J_{\mathrm{E}}(\theta) \geqslant J_{\mathrm{B}}(\theta) \quad (17b)$$

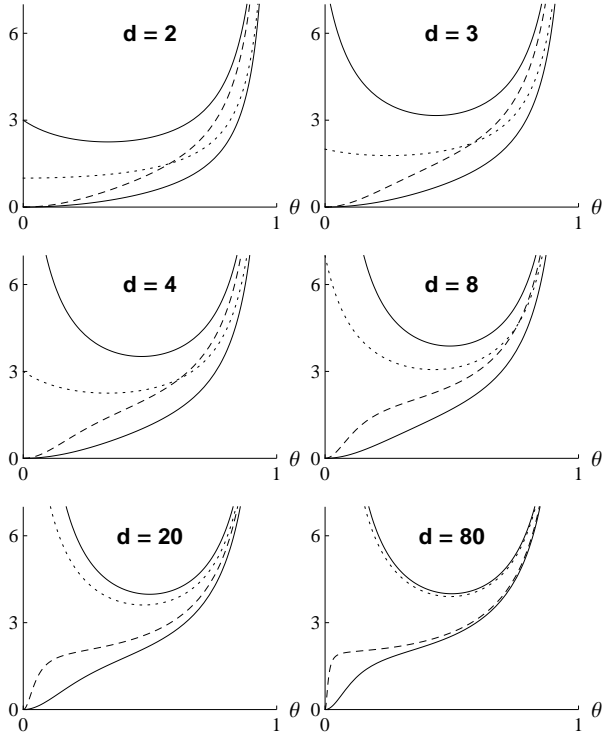for all θ and all channel dimensions d, as shown in Fig. 2.

FIG. 2. Quantum Fisher information per channel use: Scheme E (upper solid line), Scheme O (dotted), Scheme B (dashed), Scheme O,2 (solid lower).
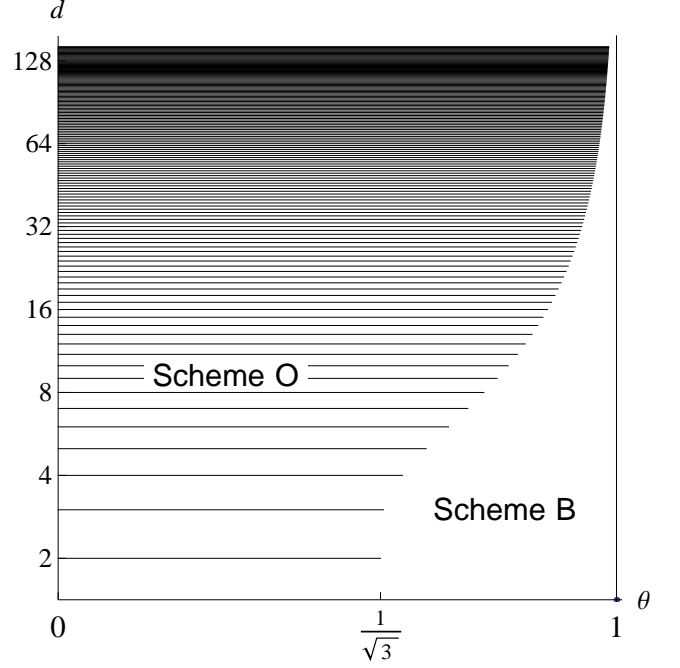

FIG. 3. Unentangled probes (scheme O) versus maximally entangled probe pairs (scheme B). Scheme B yields more information above a dimension-dependent threshold.

Comparing just schemes O and B, Fig. 3 shows that, above a d-dependent threshold for θ, scheme B always yields more information (per channel use) than does scheme O and that below $\theta = 1/\sqrt{3}$ scheme O always dominates scheme B.

Ancilla depolarization: Identification scheme E uniformly dominates schemes O and B for the qudit depolarizing channel. Scheme E, though, may be unrealistic in assuming that the ancilla qudit in the entangled pair undergoes no degradation. Suppose we model the evolution of the ancilla qudit by a depolarizing channel Γ_η with depolarizing probability $1 - \eta$. If the probe and its ancilla qudit are initially in a maximally entangled state $\sigma = |\mu\rangle\langle\mu|$ (see Eq. (11)), the channel output $\rho_\theta = (\Gamma_\eta \otimes \Gamma_\theta)(\sigma)$ is

$$\rho_\theta = \frac{1 - \eta\theta}{d^2}\mathbf{I}_{d^2} + \eta\theta\sigma. \qquad (18)$$

This is identically the parametric family of states that would result from passing a d^2-dimensional qudit in a pure state σ with no external entanglement through a depolarizing channel with parameterization $h(\theta) = \eta\theta$. Therefore, from (9), the quantum Fisher information per use of Γ_θ with this scheme E_η is

$$J_{E_\eta}(\theta) = \frac{\eta^2}{(1 - \eta\theta)\left(\eta\theta + \frac{1}{d^2 - 1}\right)}. \qquad (19)$$

Unlike identification scheme $E = E_1$, scheme E_η for $\eta < 1$ does not uniformly dominate schemes O and B. In fact, for any degree of depolarization $\eta < 1$ in the ancilla system, there exist θ where $J_{E_\eta}(\theta) < J_O(\theta)$ and/or $J_{E_\eta}(\theta) < J_B(\theta)$. Simple calculation shows, for example, that $J_{E_\eta}(\theta) < J_O(\theta)$ if $\theta > g(\eta)$ where

$$g(\eta) = \frac{\eta^2(d + 1) - 1}{\eta(1 - \eta)(d^2 - 2) + \eta^2 d}. \qquad (20)$$

Since $g(\eta) < 1$ for any $\eta < 1$ and any d, there is always a θ large enough such that $J_{E_\eta}(\theta) < J_O(\theta)$. Similar statements can be made regarding $J_{E_\eta}(\theta)$ relative to $J_B(\theta)$. The broad point here is that, in this as in many other situations, realizing the theoretical advantage of entanglement depends on precise control of the experimental apparatus, and particularly in this case, evolution of the ancilla qudit.

Probe re-circulation: One might ask whether any advantage can accrue to circulating the probe back through the channel, this being the scheme where the probe is passed through the channel, kept intact (no measurement), passed through the channel again, and, finally, measured. Conceivably, the probe could be re-circulated through the channel any number of times before measurement. Our schemes O, E, and B all accommodate probe re-circulation. The qudit depolarizing channel satisfies $\Gamma_\theta^{\circ n} = \Gamma_{\theta^n}$ where n is the number of channel passages so,

from Eq. (9),

$$J_{\mathrm{O,n}}(\theta) = \frac{n\theta^{2(n-1)}}{(1-\theta^n)\left(\theta^n + \frac{1}{d-1}\right)} \qquad (21\mathrm{a})$$

$$J_{\mathrm{E,n}}(\theta) = \frac{n\theta^{2(n-1)}}{(1-\theta^n)\left(\theta^n + \frac{1}{d^2-1}\right)} \qquad (21\mathrm{b})$$

$$J_{\mathrm{B,n}}(\theta) = \frac{2n\theta^{2(2n-1)}}{(1-\theta^{2n})\left(\theta^{2n} + \frac{1}{d^2-1}\right)} \qquad (21\mathrm{c})$$

where $J_{\mathrm{O,n}}(\theta)$, $J_{\mathrm{E,n}}(\theta)$ and $J_{\mathrm{B,n}}(\theta)$ are, respectively, the quantum Fisher informations per channel use (n channel uses with schemes O and E, $2n$ uses with B) about θ available by schemes O, E and B by measurement after n circulations of the probe through the channel. Some elementary calculus shows that $J_{\mathrm{O,n+1}}(\theta) \leqslant J_{\mathrm{O,n}}(\theta)$ for all $n \geqslant 1$, $d \geqslant 2$ and $0 \leqslant \theta \leqslant 1$, meaning that no degree of probe re-circulation is advantageous with scheme O. Since $J_{\mathrm{E,n}}(\theta;d) = J_{\mathrm{O,n}}(\theta;d^2)$ and $J_{\mathrm{B,n}}(\theta;d) = J_{\mathrm{O,2n}}(\theta;d^2)$, the same must be true for schemes E and B. The information $J_{\mathrm{O,2}}(\theta)$ with scheme O,2 is shown in Fig. 2 for various channel dimensions d. We see that with increasing dimension the entanglement schemes E and B yield information increasingly close to schemes O,1 and O,2 without entanglement: $J_{\mathrm{E}}(\theta) \to J_{\mathrm{O,1}}(\theta)$ and $J_{\mathrm{B}}(\theta) \to J_{\mathrm{O,2}}(\theta)$.

We observed that probe re-circulation, with any of our schemes, can only diminish the quantum Fisher information in the channel output. These observations were made, though, on the basis of quantum Fisher information per channel use. If channel use is not a significant cost, probe re-circulation can be advantageous. In fact, when the channel depolarizing probability $1-\theta$ is small, re-circulating the probe just once can yield more information than entanglement. For the qubit depolarizing channel with $\theta = .9$, for example, the quantum Fisher information gained by measuring a once re-circulated scheme O probe is $2J_{\mathrm{O,2}}(.9;2) = 9.42$, while that gained by measuring a maximally entangled probe with no re-circulation is $J_{\mathrm{E,1}}(.9;2) = 8.11$. Entanglement clearly holds promise for channel probing, but depending on the relative cost of this resource, probe re-circulation or some other similarly simple classical strategy may be more effective. In the case of the depolarizing channel, since probe re-circulation may be, for example, just a matter of maintaining the probe in the target depolarizing medium for some longer period of time, entanglement may need to be very inexpensive to be competitive.

IV. PROBING WITH PARTIAL ENTANGLEMENT

Various quantum Fisher informations $J(\theta)$ in the output of the depolarizing channel probed by qudits with zero and maximal entanglement were presented and compared in the previous section. Here we consider qudit probes just partially entangled with an ancilla. The effect of partial entanglement is not obvious; channels are known for which a partially entangled probe can yield more information than a fully entangled probe [4]. The essential conclusion of this section is that, for the qudit depolarizing channel of any dimension, partial entanglement always yields information intermediate between $J_{\mathrm{O}}(\theta)$ from an unentangled qudit probe and $J_{\mathrm{E}}(\theta)$ from a maximally entangled probe.

The previous section's approach, represented by use of Eq. (9), is applicable to certain special kinds of partial entanglement—qudits which are fully entangled on just a subspace. We briefly pursue this point and then turn to the question of partial entanglement generally. Consider a pair of qudits in the partially entangled pure state $\sigma = |\Psi\rangle\langle\Psi|$ with Schmidt decomposition

$$|\Psi\rangle = \psi_1|11\rangle + \ldots + \psi_d|dd\rangle \qquad (22)$$

where $\psi_j \geqslant 0$ and $\psi_1^2 + \cdots + \psi_d^2 = 1$. Suppose that d_o of the ψ_i are $1/\sqrt{d_o}$ and that the remainder are zero, where $2 < d_o < d$. These qudits are maximally entangled on a subspace of dimension $d_o < d$ with no entanglement outside this subspace. The output of the depolarizing channel probed by the first of these qudits is

$$\rho_\theta = \frac{1-\theta}{dd_o}\mathbf{I}_d \otimes \mathbf{I}_{d_o} + \theta\sigma \qquad (23)$$

and associated with this output is the quantum Fisher information

$$J(\theta) = \frac{1}{(1-\theta)\left(\theta + \frac{1}{dd_o-1}\right)}. \qquad (24)$$

This shows for qudits with this special form of partial entanglement that

$$J_{\mathrm{O}}(\theta) \leqslant J(\theta) \leqslant J_{\mathrm{E}}(\theta). \qquad (25)$$

We now prove that Eq. (25) is true generally for any partial entanglement.

Suppose two qudits are partially entangled generally as in Eq. (22). Probing the channel with the first qudit produces the channel output

$$\rho_\theta = (\Gamma_\theta \otimes \mathbf{I})(\sigma) = \frac{1-\theta}{d}\,\mathbf{I}_d \otimes \mathbf{D} + \theta\sigma \qquad (26)$$

where

$$\mathbf{D} = \sum_{i=1}^{d} \psi_i^2|i\rangle\langle i| \qquad (27)$$

operates on a single qudit. Redefining $\mathbf{D} := \sum_{i=1}^{d} \psi_i^2|ii\rangle\langle ii|$ as an operator on both qudits gives $\rho_\theta = \rho_C + \rho_Q$ where

$$\rho_Q = \frac{1-\theta}{d}\mathbf{D} + \theta|\Psi\rangle\langle\Psi| \quad \text{and} \quad \rho_C = \frac{1-\theta}{d}\mathbf{C} \qquad (28)$$

are positive semi-definite operators with orthogonal supports $\mathcal{S} = \{|ii\rangle, i = 1, \ldots, d\}$ and $\mathcal{S}^\perp = \{|ij\rangle, i \neq j\}$, respectively, determined by the Schmidt decomposition Eq. (22), and $\mathbf{C} = \sum_{j \neq i} \psi_j^2 |ij\rangle\langle ij|$.

A general result, whose proof is straightforward, is that, if $\rho_\theta = \rho_{1\theta} + \rho_{2\theta}$ where $\rho_{1\theta}$ and $\rho_{2\theta}$ have orthogonal supports, then the score operator associated with ρ_θ, i.e. $\mathbf{L}$, can be expressed as $\mathbf{L} = \mathbf{L}_1 + \mathbf{L}_2$ where $\mathbf{L}_i$ is the score operator associated with $\rho_{i\theta}$. Also the support of $\mathbf{L}_i$ is orthogonal to that of $\mathbf{L}_j$ and $\rho_{j\theta}$ if $i \neq j$. Thus $\mathbf{L} = \mathbf{L}_Q + \mathbf{L}_C$ where $\mathbf{L}_Q \rho_C = \mathbf{L}_C \rho_Q = 0$ such that $\mathbf{L}_Q$ and $\mathbf{L}_C$ satisfy, respectively,

$$\mathbf{L}_Q \rho_Q + \rho_Q \mathbf{L}_Q = 2\partial_\theta \rho_Q = 2\left(|\Psi\rangle\langle\Psi| - \frac{1}{d}\mathbf{D}\right), \quad (29)$$

$$\mathbf{L}_C \rho_C + \rho_C \mathbf{L}_C = 2\partial_\theta \rho_C = -\frac{2}{d}\mathbf{C}. \quad (30)$$

We have immediately from Eq. (30), then, that

$$\mathbf{L}_C = -\frac{1}{1-\theta}\mathbf{I}_C \quad (31)$$

where $\mathbf{I}_C$ is the identity on $\mathcal{S}^\perp$. While no similarly simple solution exists for Eq. (29), substituting ρ_Q from Eq. (28) in Eq. (29) yields

$$\langle\Psi|\mathbf{L}_Q|\Psi\rangle = \frac{d-1}{d\theta} - \frac{1-\theta}{d\theta}\operatorname{Tr}(\mathbf{D}\mathbf{L}_Q). \quad (32)$$

We now derive, using our results from Eqs. (29) and (30), an expression for the quantum Fisher information $J(\theta)$ in the channel output Eq. (26),

$$\begin{aligned}
J(\theta) &= \operatorname{Tr}(\mathbf{L}\partial_\theta \rho_\theta) \\
&= \operatorname{Tr}(\mathbf{L}_C \partial_\theta \rho_C) + \operatorname{Tr}(\mathbf{L}_Q \partial_\theta \rho_Q) \\
&= \operatorname{Tr}\left[\mathbf{L}_C\left(-\frac{1}{d}\mathbf{C}\right)\right] + \operatorname{Tr}\left[\mathbf{L}_Q\left(|\Psi\rangle\langle\Psi| - \frac{1}{d}\mathbf{D}\right)\right] \\
&= \frac{d-1}{d(1-\theta)} + \langle\Psi|\mathbf{L}_Q|\Psi\rangle - \frac{1}{d}\operatorname{Tr}(\mathbf{D}\mathbf{L}_Q) \\
&= \frac{d-1}{d(1-\theta)} + \frac{d-1}{d\theta} - \frac{1-\theta}{d\theta}\operatorname{Tr}(\mathbf{D}\mathbf{L}_Q) \\
&\quad - \frac{1}{d}\operatorname{Tr}(\mathbf{D}\mathbf{L}_Q) \\
&= \frac{d-1}{d\theta(1-\theta)} - \frac{1}{d\theta}\operatorname{Tr}(\mathbf{D}\mathbf{L}_Q). \\
&= \frac{d-1}{d\theta(1-\theta)} - \frac{1}{d\theta}\mathbf{1}^\mathsf{T}\mathbf{D}\mathbf{x} \quad (33)
\end{aligned}$$

where $\mathbf{x}$ is the d dimensional vector

$$\mathbf{x} = \begin{pmatrix} L_{11} \\ L_{22} \\ \vdots \\ L_{dd} \end{pmatrix} \quad (34)$$

with $L_{ii} = \langle ii|\mathbf{L}_Q|ii\rangle$, $\mathbf{D}$ is, now, a $d \times d$ diagonal matrix with diagonal entries $\psi_1^2, \ldots, \psi_d^2$, and $\mathbf{1}$ is the d dimensional column vector of ones. Expression (33) shows, significantly, that only $\mathbf{L}_Q$'s diagonal elements are needed

for $J(\theta)$. In appendix A we demonstrate that

$$\mathbf{1}^\mathsf{T}\mathbf{D}\mathbf{x} = \frac{d-1}{1-\theta+d\theta} - \frac{d^2\theta}{(1-\theta)(1-\theta+d\theta)^2}\mathbf{b}^\mathsf{T}\mathbf{Z}^{-1}\mathbf{b} \quad (35)$$

where $\mathbf{b}$ is a $d(d-1)/2$ dimensional vector with lexicographically ordered entries

$$\mathbf{b} = 2\begin{pmatrix} \psi_1\psi_2 \\ \psi_1\psi_3 \\ \vdots \\ \psi_{d-1}\psi_d \end{pmatrix} \quad (36)$$

and $\mathbf{Z}$ is the $d(d-1)/2 \times d(d-1)/2$ matrix

$$\mathbf{Z} = \mathbf{J} + \alpha\boldsymbol{\Omega}^\mathsf{T}\boldsymbol{\Omega} \quad (37)$$

where $\alpha = d\theta/(1-\theta+d\theta)$. The two constituent matrices of $\mathbf{Z}$ are

$$\mathbf{J} = \begin{pmatrix} \psi_1^2 + \psi_2^2 & 0 & \cdots & 0 \\ 0 & \psi_1^2 + \psi_3^2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \psi_{d-1}^2 + \psi_d^2 \end{pmatrix} \quad (38)$$

and a $d \times d(d-1)/2$ matrix $\boldsymbol{\Omega}_d = f_d(\psi_1, \ldots, \psi_d)$ with $f_d(\cdot)$ is defined recursively by

$$f_2(x, r) = \begin{pmatrix} r \\ x \end{pmatrix} \quad \text{and} \quad (39a)$$

$$f_{d+1}(x, \mathbf{r}_d) = \begin{pmatrix} \mathbf{r}_d & \mathbf{0} \\ x\mathbf{I}_d & f_d(\mathbf{r}_d) \end{pmatrix} \quad (39b)$$

where $\mathbf{r}_d = (r_1 \ldots r_d)$ and $\mathbf{0}$ is a row vector of $(d-1)(d-2)/2$ zeros. For example

$$\begin{aligned}
\boldsymbol{\Omega}_4 &= f_4(\psi_1, \psi_2, \psi_3, \psi_4) \\
&= \begin{pmatrix} \psi_2 & \psi_3 & \psi_4 & 0 & 0 & 0 \\ \psi_1 & 0 & 0 & \psi_3 & \psi_4 & 0 \\ 0 & \psi_1 & 0 & \psi_2 & 0 & \psi_4 \\ 0 & 0 & \psi_1 & 0 & \psi_2 & \psi_3 \end{pmatrix}. \quad (40)
\end{aligned}$$

Combining Eqs. (33) and (35) yields

$$\begin{aligned}
J(\theta) &= \frac{1}{(1-\theta)(\theta + \frac{1}{d-1})} \\
&\quad + \frac{d}{(1-\theta)(1-\theta+d\theta)^2}\mathbf{b}^\mathsf{T}\mathbf{Z}^{-1}\mathbf{b}. \quad (41)
\end{aligned}$$

While $\mathbf{Z}$ is singular in cases where the qudits' entanglement is confined to a proper subspace, $\mathbf{b}^\mathsf{T}\mathbf{Z}^{-1}\mathbf{b}$ is always well-defined. With this in mind one can readily check that, for maximal entanglement on a subspace, Eq. (41) is identically Eq. (24).

To prove Eq. (25) we must prove that

$$\mathbf{b}^\mathsf{T}\mathbf{Z}^{-1}\mathbf{b} \leqslant (d-1)\frac{1-\theta+d\theta}{1-\theta+d^2\theta} \quad (42)$$

or, equivalently, that

$$\mathbf{b}^{\mathsf{T}}(\alpha\boldsymbol{\Omega}^{\mathsf{T}}\boldsymbol{\Omega} + \mathbf{J})^{-1}\mathbf{b} \leqslant \frac{1}{\alpha + \frac{1}{d-1}} \ . \qquad (43)$$

Let $\lambda_{\min}[\mathbf{M}]$ and $\lambda_{\max}[\mathbf{M}]$ denote, respectively, the minimum and the maximum eigenvalues of a positive semidefinite matrix $\mathbf{M}$, and let $\mathbf{A}_d$ denote the $d \times d$ matrix $\mathbf{A}_d = \boldsymbol{\Omega}\mathbf{J}^{-1}\boldsymbol{\Omega}^{\mathsf{T}}$. Let ψ be the length-d column vector with components ψ_j. Using the relation $\mathbf{b} = \boldsymbol{\Omega}^{\mathsf{T}}\psi$, the matrix inversion lemma [22], and the Rayleigh-Ritz theorem [22], we have

$$\begin{aligned}
\mathbf{b}^{\mathsf{T}}(\alpha\boldsymbol{\Omega}^{\mathsf{T}}\boldsymbol{\Omega} + \mathbf{J})^{-1}\mathbf{b} &= \psi^{\mathsf{T}}\boldsymbol{\Omega}(\alpha\boldsymbol{\Omega}^{\mathsf{T}}\boldsymbol{\Omega} + \mathbf{J})^{-1}\boldsymbol{\Omega}^{\mathsf{T}}\psi \\
&= \psi^{\mathsf{T}}\left(\alpha\mathbf{I}_d + \mathbf{A}_d^{-1}\right)^{-1}\psi \\
&\leqslant \lambda_{\max}\left[\left(\alpha\mathbf{I}_d + \mathbf{A}_d^{-1}\right)^{-1}\right]\max_{\psi}\left(\psi^{\mathsf{T}}\psi\right) \\
&= \lambda_{\min}^{-1}\left[\alpha\mathbf{I}_d + \mathbf{A}_d^{-1}\right] \\
&= \frac{1}{\alpha + \lambda_{\min}[\mathbf{A}_d^{-1}]} \\
&= \frac{1}{\alpha + \lambda_{\max}^{-1}[\mathbf{A}_d]}. \qquad (44)
\end{aligned}$$

According to Eqs. (43) and (44), we only need to show that $\lambda_{\max}[\mathbf{A}_d] = d-1$ to prove the inequality of Eq. (25). First we show that $d-1$ is an eigenvalue of $\mathbf{A}_d$; then we show that it is the maximum eigenvalue. Let $\boldsymbol{\Lambda}$ be an eigenvector of $\mathbf{A}_d$ with associated eigenvalue λ so that

$$\mathbf{A}_d\boldsymbol{\Lambda} = \lambda\boldsymbol{\Lambda}. \qquad (45)$$

Let $\boldsymbol{\Lambda}_1 = \mathbf{J}^{-1}\boldsymbol{\Omega}^{\mathsf{T}}\boldsymbol{\Lambda}$. Then

$$\mathbf{J}^{-1}\boldsymbol{\Omega}^{\mathsf{T}}\boldsymbol{\Omega}\boldsymbol{\Lambda}_1 = \lambda\boldsymbol{\Lambda}_1 \ , \qquad (46)$$

so λ is also an eigenvalue of $\mathbf{J}^{-1}\boldsymbol{\Omega}^{\mathsf{T}}\boldsymbol{\Omega}$. One can check by direct substitution that Eq. (46) is satisfied by $\lambda = d-1$ with (unnormalized) eigenvector

$$\boldsymbol{\Lambda}_1 = \left(\frac{1}{\psi_1\psi_2} \quad \frac{1}{\psi_1\psi_3} \quad \cdots \quad \frac{1}{\psi_{d-1}\psi_d}\right)^{\mathsf{T}}. \qquad (47)$$

In appendix B we prove that $d-1$ is the maximum eigenvalue of $\mathbf{A}_d$, and this completes the proof of Eq. (25).

V. OPTIMAL MEASUREMENTS

Realizing the precision promised by the quantum Fisher information in the Cramér-Rao inequality depends on optimal measurement of the channel output. The quantum measurement that saturates the inequality is not always obvious, and in some cases—certain vectors of parameters, for example—no such measurement exists [19]. For the qudit depolarizing channel, is is possible to describe a measurement which yields a (classical) Fisher information which equals the quantum Fisher information. For this channel the density operator after evolution is such that its eigenstates are independent of the

channel parameter (only the eigenvalues depend on the channel parameter), i.e.

$$\rho_\theta = \sum_i \lambda_i(\theta)|s_i\rangle\langle s_i| \qquad (48)$$

where $\lambda_i(\theta)$ and $|s_i\rangle$ are the eigenvalues and eigenstates respectively. Then [18] the quantum Fisher information is

$$J(\theta) = \sum_i \frac{1}{\lambda_i(\theta)}\left(\frac{d\lambda_i}{d\theta}\right)^2. \qquad (49)$$

However [18], this is exactly the expression for the classical Fisher information for a probability distribution with probabilities $\{\lambda_i(\theta)\}$. Thus the quantum Fisher information can be attained via a measurement which results in this probability distribution. Clearly measurement in the eigenbasis of ρ_θ accomplishes this. Here, one of the eigenstates of ρ_θ is the initial state $|\mu\rangle$ of the pair of qudits prior to action by the depolarizing channel. It follows that the projective measurement for which the projectors are

$$\Pi_0 = |\mu\rangle\langle\mu| \quad \text{and} \qquad (50)$$
$$\Pi_1 = I - |\mu\rangle\langle\mu| \qquad (51)$$

yields a probability distribution, which gives a classical Fisher information exactly equal to that of the quantum Fisher information. For an entangled input state such as that used in scheme E, this optimal measurement requires measurements involving entangled states. It is unclear whether there is a scheme which involves unentangled states at the measurement phase and which attains the quantum Fisher information; such a scheme exists for unitary parameter estimation [23].

VI. SUMMARY

We have given an extensive treatment of channel identification for the general qudit depolarizing channel, comparing different probe preparation schemes based on the quantum Fisher information attainable by channel probing. We showed for depolarizing channels of any dimension that maximally entangling the channel probe with an external qubit uniformly realizes the maximum theoretical advantage, but that this advantage disappears if the ancilla qubit is not sufficiently shielded from depolarization. We primarily compared the same three channel identification schemes considered in [1, 2], though our results apply generally for dimension $d < \infty$. We departed from [1, 2] most significantly in choosing to compare identification schemes on an information per channel use basis. This is a practically meaningful basis of comparison; in any case, our results for each scheme are readily amenable to other comparisons.

We have considered here identification schemes involving bipartite probe entanglement, with an ancilla qudit

or with another probe. Multipartite probe entanglement of various orders and types remains for study. We hope to be able to address at least some important multipartite entanglement schemes with our present methods.

Appendix A: Diagonal elements of the score operator

Substituting from Eq. (28) into Eq. (29) gives

$$-\frac{1}{d}\mathbf{D} + |\Psi\rangle\langle\Psi| = \frac{1-\theta}{2d}\left(\mathbf{L}_Q\mathbf{D} + \mathbf{D}\mathbf{L}_Q\right)$$
$$+\frac{\theta}{2}\left(\mathbf{L}_Q|\Psi\rangle\langle\Psi|+|\Psi\rangle\langle\Psi|\mathbf{L}_Q\right).\text{(A1)}$$

Acting with this on $|jj\rangle$ and acting on the result with $\langle ii|$ gives

$$-\frac{2}{d}\psi_i^2\delta_{ij} + 2\psi_i\psi_j = \frac{1-\theta}{d}\left(\psi_i^2 + \psi_j^2\right)L_{ij}$$
$$+\theta\sum_{k=1}^{d}\left(\psi_j\psi_k L_{ik} + \psi_i\psi_k L_{jk}\right)(\text{A2})$$

where $L_{ij} = \langle ii|\mathbf{L}_Q|jj\rangle$. Note that the score operator is Hermitian and in this case real; thus $L_{ij} = L_{ji}$. Thus Eq. (A2) is a system of $d(d+1)/2$ distinct linear equations for the components of $\mathbf{L}_Q$. Our aim is to invert these for the diagonal elements of $\mathbf{L}_Q$. Using $\mathbf{x}$ as defined in Eq. (34) and

$$\mathbf{y} = \begin{pmatrix} L_{12} \\ L_{13} \\ \vdots \\ L_{1d} \\ L_{23} \\ \vdots \\ L_{d-1d} \end{pmatrix} \tag{A3}$$

gives

$$\begin{pmatrix} \mathbf{R} & \mathbf{S} \\ \mathbf{S}^{\mathsf{T}} & \mathbf{T} \end{pmatrix} \begin{pmatrix} \mathbf{x} \\ \mathbf{y} \end{pmatrix} = \begin{pmatrix} \mathbf{a} \\ \mathbf{b} \end{pmatrix} \tag{A4}$$

where $\mathbf{a}$ is a d dimensional vector given by

$$\mathbf{a} = \frac{d-1}{d}\begin{pmatrix} \psi_1^2 \\ \psi_2^2 \\ \vdots \\ \psi_d^2 \end{pmatrix}, \tag{A5}$$

$\mathbf{b}$ is given in Eq. (36), and the concatenation $\left(\mathbf{x}\ \mathbf{y}\right)^{\mathsf{T}}$ is the column vector consisting of the components of $\mathbf{x}$ followed by those of $\mathbf{y}$. The submatrices $\mathbf{R}$ and $\mathbf{S}$ within Eq. (A4) are readily derived from Eq. (A2) and are

$$\mathbf{R} = \begin{pmatrix} \rho_{11} & 0 & \cdots & 0 \\ 0 & \rho_{22} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \rho_{dd} \end{pmatrix} = \frac{1-\theta+d\theta}{d}\mathbf{D} \tag{A6}$$

and

$$\mathbf{S} = \theta\mathbf{D}^{1/2}\mathbf{\Omega}_d \tag{A7}$$

where $\mathbf{\Omega}_d$ is as defined via Eq. (39b). The final submatrix is

$$\mathbf{T} = \frac{1-\theta}{d}\mathbf{J} + \theta\mathbf{\Omega}_d^{\mathsf{T}}\mathbf{\Omega}_d \tag{A8}$$

and this can be verified by considering Eq. (A2) for $i < j$. The left side yields the row of $\mathbf{b}$ corresponding to the indices i and j and the right side yields the following contribution from $\mathbf{Ty}$

$$\frac{1-\theta}{d}\left(\psi_i^2 + \psi_j^2\right)L_{ij}$$
$$+\theta\left(\sum_{k=1,k\neq i}^{d}\psi_j\psi_k L_{ik} + \sum_{k=1,k\neq j}^{d}\psi_i\psi_k L_{jk}\right). \tag{A9}$$

The first term clearly corresponds to $(1-\theta)/d$ $\mathbf{Jy}$. To check the second term, we can show by an iterative process that

$$\mathbf{\Omega}_d\mathbf{y} = \begin{pmatrix} \sum_{k\neq 1}\psi_k L_{1k} \\ \vdots \\ \sum_{k\neq d}\psi_k L_{dk} \end{pmatrix}. \tag{A10}$$

Using this, we can again show iteratively that

$$\mathbf{\Omega}_d^{\mathsf{T}}\mathbf{\Omega}_d\mathbf{y} = \begin{pmatrix} \sum_{k\neq 1}\psi_k L_{1k} + \sum_{k\neq 2}\psi_k L_{2k} \\ \sum_{k\neq 1}\psi_k L_{1k} + \sum_{k\neq 3}\psi_k L_{3k} \\ \vdots \\ \sum_{k\neq d-1}\psi_k L_{d-1k} + \sum_{k\neq d}\psi_k L_{dk} \end{pmatrix} \tag{A11}$$

and this corresponds to the second term. This justifies the form of Eq. (A4).

Applying the matrix inversion lemma [22] to the solution of Eq. (A4) for $\mathbf{x}$ gives

$$\mathbf{x} = \left(\mathbf{R} - \mathbf{ST}^{-1}\mathbf{S}^{\mathsf{T}}\right)^{-1}\left(\mathbf{a} - \mathbf{ST}^{-1}\mathbf{b}\right)$$
$$= \left[\mathbf{R}^{-1} + \mathbf{R}^{-1}\mathbf{S}(\mathbf{T} - \mathbf{S}^{\mathsf{T}}\mathbf{R}^{-1}\mathbf{S})^{-1}\mathbf{S}^{\mathsf{T}}\mathbf{R}^{-1}\right]$$
$$(\mathbf{a} - \mathbf{ST}^{-1}\mathbf{b}) \tag{A12}$$

From (A12), and using

$$\mathbf{R}^{-1}\mathbf{a} = \frac{d-1}{1-\theta+d\theta}\mathbf{1} \tag{A13}$$

a lengthy calculation yields

$$\mathbf{Dx} = \frac{d}{1-\theta+d\theta}(\mathbf{a} - \mathbf{ST}^{-1}\mathbf{b})$$
$$+ \frac{d}{1-\theta+d\theta}\mathbf{S}(\mathbf{T} - \mathbf{S}^{\mathsf{T}}\mathbf{R}^{-1}\mathbf{S})^{-1}\mathbf{S}^{\mathsf{T}}$$
$$\left[\frac{d-1}{1-\theta+d\theta}\mathbf{1} - \mathbf{R}^{-1}\mathbf{ST}^{-1}\mathbf{b}\right]. \tag{A14}$$

Using the relations $\mathbf{1}^\mathsf{T}\mathbf{a} = (d-1)/d$ and $\mathbf{S}^\mathsf{T}\mathbf{1} = \theta\mathbf{b}$, we calculate from Eq. (A14) that

$$\mathbf{1}^\mathsf{T}\mathbf{D}\mathbf{x} = \frac{d-1}{1-\theta+d\theta} + \frac{d\theta}{(1-\theta+d\theta)^2}$$
$$\times \mathbf{b}^\mathsf{T}(\mathbf{S}^\mathsf{T}\mathbf{R}^{-1}\mathbf{S} - \mathbf{T})^{-1}\mathbf{b}. \qquad (A15)$$

Recalling the definitions of $\mathbf{S}$ and $\mathbf{T}$ in Eqs. (A7) and (A8), we calculate further that

$$\mathbf{1}^\mathsf{T}\mathbf{D}\mathbf{x} = \frac{d-1}{1-\theta+d\theta}$$
$$- \frac{d^2\theta}{(1-\theta)(1-\theta+d\theta)^2}\mathbf{b}^\mathsf{T}\mathbf{Z}^{-1}\mathbf{b} \qquad (A16)$$

where

$$\mathbf{Z} = \mathbf{J} + \alpha\mathbf{\Omega}^\mathsf{T}\mathbf{\Omega} \qquad (A17)$$

with $\alpha = d\theta/(1-\theta+d\theta)$. This gives Eq. (35).

Appendix B: Maximum eigenvalue of $\mathbf{A}_d$

The matrix $\mathbf{A}_d = [a_{ij}]$ has elements

$$a_{ij} = \begin{cases} \sum_{k=1,k\neq i}^{d} \dfrac{\psi_k^2}{\psi_i^2 + \psi_k^2} & \text{if } i = j, \\ \dfrac{\psi_i\psi_j}{\psi_i^2 + \psi_j^2} & \text{if } i \neq j. \end{cases} \qquad (B1)$$

For example,

$$\mathbf{A}_3 = \begin{pmatrix} \frac{\psi_2^2}{\psi_1^2+\psi_2^2} + \frac{\psi_3^2}{\psi_1^2+\psi_3^2} & \frac{\psi_1\psi_2}{\psi_1^2+\psi_2^2} & \frac{\psi_1\psi_3}{\psi_1^2+\psi_3^2} \\ \frac{\psi_2\psi_1}{\psi_2^2+\psi_1^2} & \frac{\psi_1^2}{\psi_2^2+\psi_1^2} + \frac{\psi_3^2}{\psi_2^2+\psi_3^2} & \frac{\psi_2\psi_3}{\psi_2^2+\psi_3^2} \\ \frac{\psi_1\psi_3}{\psi_1^2+\psi_3^2} & \frac{\psi_2\psi_3}{\psi_2^2+\psi_3^2} & \frac{\psi_1^2}{\psi_3^2+\psi_1^2} + \frac{\psi_2^2}{\psi_3^2+\psi_2^2} \end{pmatrix}. \qquad (B2)$$

Define the angles ϕ_{ij} for $1 \leqslant i < j \leqslant d$ by $\cos\phi_{ij} = \psi_i/\sqrt{\psi_i^2 + \psi_j^2}$ and let Φ_d be the length-$d(d-1)/2$ vector $\Phi_d = (\phi_{12} \quad \phi_{13} \quad \ldots \quad \phi_{d-1,d})$. The matrix $\mathbf{A}_d$ is a function $\mathbf{A}_d(\Phi_d)$ of the angles in Φ_d. Consider the characteristic function $h(\lambda) = \det(\lambda\mathbf{I}_d - \mathbf{A}_d)$ of $\mathbf{A}_d$. The eigenvalue $\lambda_{\max}[\mathbf{A}_d] = d - 1$ if and only if $h'(\lambda)$ is positive for $\lambda > d - 1$. This is a consequence of the following. If the eigenvalues of $\mathbf{A}_d$ are $\lambda_{\min}, \ldots, \lambda_{\max}$ then $h(\lambda) = (\lambda - \lambda_{\min})\ldots(\lambda - \lambda_{\max})$ and $h'(\lambda) = h(\lambda)[1/(\lambda - \lambda_{\min}) + \ldots 1/(\lambda - \lambda_{\max})]$. Clearly if $\lambda > \lambda_{\max}$ then $h(\lambda) > 0$ and $h'(\lambda) > 0$. A graphical representation of the characteristic function reveals that at some point between $\lambda_{\max}$ and the eigenvalue immediately smaller than this, $h'(\lambda)$ must be negative since $h(\lambda)$ in this region is between two of its roots. Thus if $h'(\lambda) > 0$ for all $\lambda > \tilde{\lambda}$ then $\tilde{\lambda} = \lambda_{\max}$. We have

$$h'(\lambda) = \sum_{i=1}^{d} \det(\lambda\mathbf{I}_{d-1} - \mathbf{A}_d^{ii}) \qquad (B3)$$

where $\mathbf{A}_d^{ii}$ is the matrix $\mathbf{A}_d$ with its ith row and column deleted. The summand in Eq. (B3) is the characteristic

function of $\mathbf{A}_d^{ii}$, so we need just show that no $\mathbf{A}_d^{ii}$ has an eigenvalue greater than $d-1$. A standard calculation yields

$$\mathbf{A}_d^{ii} = A_{d-1}(\Phi_d^i) + \Delta^i \qquad (B4)$$

where Δ^i is a diagonal matrix with diagonal elements $\cos^2\phi_{ij}$, and Φ_d^i is the vector Φ_d with the $d-1$ angles ϕ_{ij} and ϕ_{ji} deleted. Therefore, by Weyl's inequality [22],

$$\lambda_{\max}[\mathbf{A}_d^{ii}] \leqslant \lambda_{\max}[A_{d-1}(\Phi_d^i)] + \lambda_{\max}[\Delta^i]$$
$$\leqslant \lambda_{\max}[A_{d-1}(\Phi_d^i)] + 1. \qquad (B5)$$

We have $\lambda_{\max}[\mathbf{A}_2(\Phi_2)] = \lambda_{\max}[\mathbf{A}_2(\Phi_3^i)] = 1$ by direct calculation from Eq. (B1). Thus $\lambda_{\max}[\mathbf{A}_3^{ii}] \leqslant 2$ by Eq. (B5), and $\lambda_{\max}[\mathbf{A}_3] \leqslant 2$ by consideration of Eq. (B3). Arguing recursively, we conclude that $\lambda_{\max}[\mathbf{A}_d] \leqslant d-1$ for $d \geqslant 2$. We saw that $d-1$ is an eigenvalue of $\mathbf{A}_d$ so $\lambda_{\max}[\mathbf{A}_d] = d-1$.

[1] A. Fujiwara, Phys. Rev. A **63**, 042304 (Mar 2001).
[2] M. Sasaki, M. Ban, and S. M. Barnett, Phys. Rev. A **66**, 022308 (Aug 2002).
[3] A. Fujiwara and H. Imai, J. Phys. A **36**, 8093.
[4] A. Fujiwara, Phys. Rev. A **70**, 012317 (Jul 2004).
[5] M. A. Ballester, Phys. Rev. A **69**, 022303 (Feb 2004).

[6] M. Frey, A. L. Miller, L. K. Mentch, and J. Graham, Quantum Information Processing **9** (Mar 2010).

[7] M. Frey, L. Coffey, L. K. Mentch, A. L. Miller, and S. Rubin, Int. J. Quantum Information(2010).

[8] M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information* (Cambridge University Press, Cambridge, 2000).

[9] J. H. Shapiro and S. R. Shepard, Phys. Rev. A **43**, 3795 (Apr 1991).

[10] C. King, IEEE Trans. Inf. Theory **49**, 221 (Jan 2003).

[11] A. Dragan and K. Wódkiewicz, Phys. Rev. A **71**, 012322 (Jan 2005).

[12] I. B. Slimen, O. Trabelsi, H. Rezig, R. Bouallègue, and A. Boualègue, J. Comp. Sci. **3**, 424 (2007).

[13] S. Boixo and A. Monras, Phys. Rev. Lett. **100**, 100503 (Mar 2008).

[14] D. Daems, Phys. Rev. A **76**, 012310 (Jul 2007).

[15] M. Piani and J. Watrous, Phys. Rev. Lett. **102**, 250501 (Jun 2009).

[16] C. W. Helstrom, Phys. Lett. A **25**, 101 (Jul 1967).

[17] S. L. Braunstein and C. M. Caves, Phys. Rev. Lett. **72**, 3439 (May 1994).

[18] M. Paris, Int. J. Quantum Information **9**, 125 (2009).

[19] C. Ping and L. Shunlong, Front. Math. China **2**, 359 (2007).

[20] M. Frey and D. Collins, Quantum Information and Computation VII, Proc. SPIE **7342**, 73420N (2009).

[21] A. Fujiwara and P. Algoet, Phys. Rev. A **59**, 3290 (May 1999).

[22] R. A. Horn and C. R. Johnson, *Matrix Analysis* (Cambridge University Press, Cambridge, 1985).

[23] V. Giovannetti, S. Lloyd, and L. Maccone, Phys. Rev. Lett. **96**, 010401 (Jan 2006).